REMARKS

Claims 1-18 are pending and presented for reconsideration in connection with the subject application. Claims 1, 15 and 17 are the independent claims in the application.

Claims 1-18 were rejected under 35 U.S.C. §103(a) as allegedly being unpatentable over U.S. Patent No. 6,047,257 to Dewaele.

The present application relates to hands-free command and control of dental images in a dental imaging system. In many instances, a dentist (or another dental care professional) needs to refer to one or more of a patient's plural dental images, while attending to the patient. The dental images may include, for example, intra-oral images, panoramic dental images, FOTI images and periodontic images. In current dental office practice, many (if not all) of these images are stored electronically in a storage device of a dental imaging computer system. Conventional dental imaging systems typically require manual operation of computer input devices in order to specify and cause the specified dental image to be retrieved and displayed.

As discussed in the application at, for example, page 5, lines 19-27, computerized voice recognition in the methods and systems of this application is provided to enable a dentist or dental technician to specify, through spoken commands, dental images to be retrieved from storage in a computer system and

displayed and/or manipulated, without requiring the dentist/technician to manually operate computer input devices.

In addition, as discussed at the application at page 4, lines 9-26, a user's voice commands can be processed through a voice interface for user selection of options for image processing of the retrieved image. For example, a user can command the system to manipulate the image, such as rotate, resize (e.g., increase image size to full screen, increase or decrease image size by a specified percentage, etc.) or move the image on the display, pan or zoom the image, change the brightness, contrast, color preferences or other color processing settings of the image, select a region of the image for manipulation, etc.

Thus, the dental professional can cause the desired dental images to be retrieved, displayed and manipulated, while continuing to use her/his hands for attending to the patient, without risking contamination from manually operating computer input devices.

For example, independent claim 17 of the present application relates to a method of hands-free command and control of a dental imaging system. The method includes converting to electronic speech data a voice command from a user, to select for viewing one of a plurality of dental images stored in a storage device for a selected dental patient, and processing the electronic

command data to cause the selected dental image to be retrieved from the storage device and then displayed on a display monitor. Similar features are recited in independent claims 1 and 15. The claimed invention recited in dependent claims provide in addition for further processing and image manipulation of the retrieved image.

Applicants maintain that Dewaele does not disclose or suggest the invention claimed in the present application, for at least the following reasons.

Dewaele, as understood by Applicants, relates to providing identification information to be associated with an image on a photo-stimulable phosphor screen, during an image **recording** process.

The term "identification information" or "identification data" are defined in Dewaele to be data identifying a patient to which a medical image pertains, data identifying the examination type that is performed or is going to be performed, and other data that are commonly associated with a medical image, such as the name of the radiologist, the sex of the patient, etc. (see Dewaele, col. 1, lines 14-21, and col. 7, lines 23-26). Conventional systems for entry of identification information through use of speech recognition, such as proposed by Dewaele, are disclosed at page 2, lines 3-10 of the application.

Dewaele discloses that identification information may be

entered through a microphone and speech recognition to expedite the data entry process while reducing the probability of data entry error. Dewaele does not disclose or suggest, however, retrieval for display and image manipulation of a computer-stored dental (or medical) image specified through voice commands which are processed through speech recognition.

The January 28, 2003 Office Action cites Dewaele, Figure 1, elements 4, 6, 8 and 7, column 5, lines 1-65, and column 7, lines 45-55, as purported support that Dewaele teaches processing a voice command received through a microphone to select one of the plurality of images for viewing. Applicants respectfully disagree.

Figure 1 of Dewaele shows a speech recognition/synthesis subassembly 4, an antenna 6, a cassette 7 on which a radiographic image is to be recorded, and a radiofrequency tag 8 on which identification data can be stored. None of the elements are shown or disclosed as being associated with processing voice commands to select images for viewing and image manipulation.

Dewaele discusses at col. 5, lines 1-65 the advantages of using automatic voice recognition for receiving and processing identification information. Dewaele, col. 4, line 65 through col. 6, line 9 is repeated below for the record for purposes of completeness:

A strong prejudice has existed against the application of data input via speech for identification purposes. Speech recognition is difficult primarily because of variability, which comes in different forms: (1) variability of sounds (different words, phrases or subword units), (2) transducer/channel variability. Further there is a risk of interference with background noise from extraneous speech or transient acoustic events.

In the field of medical images these prejudices have been overcome because:

(1) the number of words in a medical identification task is restricted to a vocabulary of at most 100 single and isolated words so that the variability of sounds is limited.

(2) transducer/channel variability including differences in signal characterisation is limited since the input is always via microphone, the characteristics of which are known at design stage. Thus, the voice recognition system need not be able to cope with a variety of sources.

(3) the risk of interference with background noise from extraneous speech or transient acoustic events is limited on a radiology department since the voice input is under software control of the application and is restricted to well defined time slots in the course of operation.

Significant advances in several technologies and application areas pertinent to voice processing have made feasible automatic voice recognition, such as (1) smart microphones adapting to any acoustic environment and giving optimum signal-to-noise ratio in noisy backgrounds (2) acoustic echo cancellation to provide echo-free communications (3) advances in algorithms and DSP implementation of these algorithms providing high performance on reasonable cost platform. Although the sources of variability cannot be eliminated in general, speech recognition technology has reached a point to model and handle them properly. These models are based on (1) standard pattern recognition or (2) on hidden Markov models.

The first class computes a best match similarity score between a spectral pattern of features against a database of stored vocabulary patterns. These spectral patterns model differences across different speakers and variance statistics derived over the time duration of the word. The second class of models calculates the highest likelihood score for a probabilistic model for each word of a vocabulary of words.

Voice processing has proven to be <u>very well suited for the purpose of identification</u> in a hospital environment or specifically in a radiology department for the following reasons.

First, the speaking format, that is the mode of speaking to the machine has limited complexity : it will basically fall into one of the following categories:

> (a) isolated word recognition (each spoken command or data entity represents one single word) or
> (b) connected word mode (the operator uses fluent speech but with highly constrained vocabulary) or
> (c) continuous speech mode (the operator dictates phrases or performs a dialogue).
>
> The first mode is suited for control and command entry and for input of single word data, the second mode is suited for entry of letters of the alphabet or digits. The third category of speaking format is continuous speech and is applicable for voice entry of comment-like annotations or clinical protocols to a patient's identification records.
>
> A second reason why voice processing is well suited for identification of medical images is that the degree of speaker dependence is low, since the number of operators is typically low and almost fixed over time.
>
> A third reason is that the vocabulary size and complexity is low to moderate. It will typically consist of a set of command and control words to navigate the user interface of the identification application by appropriate words for operations such as screen selection, cursor movement and key stroke shortcuts. Further, it will consist of sets of words for mandatory inputs such as examination type, sub-examination type, image destination type. Finally, many identification data are letters drawn from the alphabet, or digits such as patient's birthday (digits), patient's sex (letter), patient's index (digits), number of hardcopies requested (digit), image layout parameters (letters or digit). (Emphasis added).

Although Dewaele discloses means for associating identification data with the medical image to be identified, Dewaele does not disclose or suggest, however, converting a voice command received through a microphone to electronic speech data for **selecting one of a plurality of images for viewing and image manipulation**.

Col. 7, lines 19-55 of Dewaele, which is also cited in the Office Action, states as follows:

> The described system is a digital radiography system wherein a radiographic image is recorded on a photostimulable phosphor screen. The photostimulable phosphor screen is conveyed in a

> cassette 7. The cassette is provided with a radio-frequency tag 8 in which identification data, i.e. data concerning a patient that is subjected to a radiographic examination and concerning the type of examination that is performed etc., are stored.
>
> The system comprises an identification station 1, a read out station 2 in which the image stored in the photostimulable phosphor screen is read out and digitized and wherein the digital signal representation of the radiographic image is subjected to image processing. A laser recorder 3 is provided for reproducing the read out image.
>
> The system shown in FIG. 1 can be expanded to include other stations such as a workstation for performing off-line processing on the digital representation of the radiographic image and/or for performing soft copy diagnosis. However, **since these additional components are not relevant in the context of the present invention**, they will not be described in detail.
>
> The identification station 1 consists of a personal computer (or alternatively a workstation) which is in the described embodiment connected to a network so as to provide access to a hospital information system (HIS) or a radiology information system 9 (RIS).
>
> The identification station is further equipped with a speech recognition/synthesis subassembly 4, with a dynamic microphone input 5 to provide data input via speech and a speaker 10 to provide auditive responses. An example of a suitable speech recognition subassembly is a standalone board Star 21 of Lernout and Hauspie (Belgium) with microphone speech input and, an (AD21) DSP, speech models stored in (AMD Flash) memory, RS232 connection to host, amplifier for synthesized TTS (Text to Speech), speech output, power supply. (Emphasis added)

Applicants have carefully studied the cited portions of Dewaele, and found no disclosure or suggestion, however, of converting a user voice command to electronic speech data for selecting for viewing and image manipulation dental images stored in a storage device, as provided by the claimed invention. Indeed, Dewaele teaches away from using voice control of image retrieval manipulation.

The January 28, 2003 Office Action cites Dewaele, column 9, line 39 through column 10, line 65 as alleged support that

Dewaele discloses a command and control processor for causing a selected image to be retrieved and displayed on a monitor. Dewaele, Fig. 1, elements 4 and 6-8, column 5, line through column 6, line 6, column 7, lines 19-55, column 9, line 39 through column 10, line 65 are cited in the Office Action as alleged support that Dewaele discloses manipulation of images corresponding to a dental patient, through voice recognition of voice commands. Applicants respectfully disagree.

Dewaele discusses, at column 9, line 39 through column 10, line 65, operations performed at the identification station, including entry of assorted identification information, such as patient's name, examination type, sub-examination type, comments, etc. However, Dewaele simply does not disclose or suggest (a) processing electronic voice command data to cause a selected image to be retrieved from an electronic storage device and displayed on a monitor, and (b) manipulating the retrieved image according to voice commands processed through voice recognition.

Col. 9, line 39 through col. 11, line 16 of Dewaele is repeated below for the record for purposes of completeness:

> The following is a description of operations performed, along with details pertinent to the voice recognition functionality:
>
> A radiologist specific identification-screen is popped up either by sensing an operator's personal identification carrier to the read/write identification subsystem or by voice recognition of an utterance of the operator's name by the speech recognition subassembly. The database of voice patterns pertaining to the operator is made active.

The patient's name is uttered by the operator to identify the patient to the system. On correct recognition, the name is displayed in the patient name field. On false recognition, an alternative voice input is offered consisting of spelling the patient's name. During utterance of the letters of the name, the list of patients currently residing in the hospital as established during patient intake, is popped up onto the screen. The portion of the list displayed during spelling is continuously narrowed as more successive letters are recognized by the system. In addition to the patient name, the list also shows the running number of the patient in the list and the patient's birthday. At all times during spelling the name, a shortening may be obtained by uttering the digits of the running number of the patient as soon as the data searched for become displayed. Both spelling of 26 letters of the alphabet and the 10 digits is far less prone to recognition error than direct recognition of the patient's name, for reasons that the vocabulary of letters and digits has fixed size and can be specifically trained to the operator. In contrast, direct recognition of the patient's name is more difficult since the number of words is substantially large (as large as 500 e.g.) and since the voice sample of the name used as a reference template, has been recorded by a receptionist at patient intake. This person in general is different from the radiology operator, and patient name recognition thus has presented itself as a speaker independent recognition task. An acceptance qualifier completes the patient entry; a correction qualifiers offers the operator the opportunity to re-enter a name; a rub-out qualifier enables to erase letters in much the same way as the backspace key on a keyboard operates. As a fallback way of entry, the patient name may still be selected by cursor movement from the patient list or entered manually by keyboard on network failure or absence of a RIS database. The patient name is filled in in its appropriate field, and other patient related data are retrieved from the RIS database to complete fields such as sex (M/F) and birthday. Should these latter items be unavailable, voice entry of them is task of recognition of a sequence of letters and digits.

The system prompts the operator to input the examination type. The examination type is one out of a radiologist specific list of examination (such as thorax, pelvis, skull, . . .) and recognition thus belongs to the isolated word mode. The size of the examination list typically does not exceed 20. On correct recognition, the examination type is automatically entered into the appropriate field. On false recognition, a list of all examination types and a ranking number is popped up to assist the operator in selecting the examination type. Selection now is done by uttering the digits (one or two digits) of the ranking number. Alternatively, the user may use cursor movements to scroll through the list and the `enter` button to select.

The system then prompts the operator to input the sub-examination type. The sub-examination type is one out of a

> radiologist specific list of sub-examinations (e.g. `lateral`, `frontal`, . . .), pertaining to the examination type just selected. The size of the sub-examination list typically does not exceed 25 per examination, still amounting to a total number of sub-examinations as large as 500. However, knowledge of the examination type restricts the number of valid choices for the sub-examination in that sub-examination of other examination classes are not taken into consideration. This makes the recognition of the sub-examination more manageable. Analogously, on correct recognition, the sub-examination type is automatically entered into its field. On false recognition, a list of all examination types and a ranking number is popped up to assist the operator in selecting the sub-examination type by utterance of the corresponding digit sequence.
>
> Examination and sub-examination determine layout parameters as to how the image will be processed, printed and displayed (these include patient position, cassette position and exposure class). These parameters are retrieved from radiologist specific internal data buffers and are automatically filled out in their appropriate fields. Should these fields be modified, the operator issues voice commands as to the placement of the cursor in one of these fields and modifies the default entry.
>
> The system prompts the operator to input the destination type. The destination type is one out of a radiologist specific list of preferred hardcopy and softcopy devices to send the digitized image to. The list typically contains smaller than 10 items. Selection proceeds in a way similar to that of the examination and sub-examination entry. Next, the number of copies on a hardcopy unit is entered by voice.
>
> Optionally, the operator may enter comments in the `user info` field as a recorded voice stream upon issuing the request "info". Voice data is stored along with other identification data in a database.
>
> On completion of all fields on the identification screen, the system prompts the operator to write the data to the cassette identification carrier by means of the Read/Write subassembly on recognition of the action word "write" or other meaningful answers such as "OK" or "Yes".

Applicants have carefully reviewed the cited portion of Dewaele. Although Dewaele discloses entering orally-specified identification information which may be associated with a medical image, Applicants find no disclosure or suggestion in Dewaele, however, of processing electronic speech data corresponding to a

voice command selecting a dental image, to **<u>cause the selected image to be retrieved from the storage device and then displayed</u>** on a display monitor, as provided by the claimed invention. In addition, Applicants find no teaching or suggestion in Dewaele of manipulating the retrieved image according to voice commands processed through voice recognition, as provided by systems, apparatuses and methods of the present application.

If the Examiner has found in Dewaele teaching or suggestions of (1) processing voice commands for selecting a dental image, to cause the selected image to be retrieved from the storage device and then displayed, and (2) manipulating the retrieved image according to voice commands processed through voice recognition, Applicants respectfully request that a citation be provided to the specific lines in Dewaele on which such teaching or suggestion can be found.

Since Dewaele does not relate to voice-controlled image retrieval, display and manipulation, Dewaele cannot render the claimed invention obvious.

For at least the above-mentioned reasons, Applicants maintain that the claimed invention would not have been obvious at the time of the invention in view of Dewaele.

Accordingly, Applicants maintain that claims 1-18 are patentably distinct from the cited art, and earnestly solicit the allowance of claims 1-18.

If a telephone interview would be of assistance in advancing prosecution of the subject application, Applicants' undersigned attorneys invite the Examiner to telephone them at the telephone number provided below.

If a petition for an extension of time is required to make this response timely, this paper should be considered to be such a petition, and the Commissioner is authorized to charge the requisite fees to our Deposit Account No. 03-3125.

If any is required, authorization is hereby given to charge the amount of any such fee to Deposit Account No. 03-3125.

Respectfully submitted,

Paul Teng

Norman H. Zivin, Reg.No.25,385
Paul Teng, Reg.No.40,837
Attorneys for Applicants
Cooper & Dunham, LLP
1185 Avenue of the Americas
New York, New York 10036
(212) 278-0400

I hereby certify that this correspondence is being transmitted by facsimile transmission this date and is being deposited this date with the U.S. Postal Service with sufficient postage as first class mail in an envelope addressed to: Assistant Commissioner for Patents, Box AF, Washington, D.C. 20231.

Paul Teng — March 28, 2003

Paul Teng
Reg. No. 40,837

Date